



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

Can altmetrics reflect societal impact considerations?: Exploring the potential of altmetrics in the context of a sustainability science research center

Kassab, Omar ; Bornmann, Lutz ; Haunschild, Robin

Abstract: Societal impact considerations play an increasingly important role in research evaluation. In particular, in the context of publicly funded research, proposal templates commonly include sections to outline strategies for achieving broader impact. Both the assessment of the strategies and the later evaluation of their success are associated with challenges in their own right. Ever since their introduction, altmetrics have been discussed as a remedy for assessing the societal impact of research output. On the basis of data from a research center in Switzerland, this study explores their potential for this purpose. The study is based on the papers (and the corresponding metrics) published by about 200 either accepted or rejected applicants for funding by the Competence Center Environment and Sustainability (CCES). The results of the study seem to indicate that altmetrics are not suitable for reflecting the societal impact of research that was considered: The metrics do not correlate with the ex ante considerations of an expert panel.

DOI: https://doi.org/10.1162/qss_a_00032

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-191687>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Kassab, Omar; Bornmann, Lutz; Haunschild, Robin (2020). Can altmetrics reflect societal impact considerations?: Exploring the potential of altmetrics in the context of a sustainability science research center. *Quantitative Science Studies*, 1(2):792-809.

DOI: https://doi.org/10.1162/qss_a_00032



an open access  journal



Citation: Kassab, O., Bornmann, L., & Haunschild, R. (2020). Can altmetrics reflect societal impact considerations?: Exploring the potential of altmetrics in the context of a sustainability science research center. *Quantitative Science Studies*, 1(2), 792–809. https://doi.org/10.1162/qss_a_00032

DOI:
https://doi.org/10.1162/qss_a_00032

Received: 29 August 2019
Accepted: 08 February 2020

Corresponding Author:
Omar Kassab
omar.kassab@gess.ethz.ch

Handling Editor:
Ludo Waltman

Copyright: © 2020 Omar Kassab, Lutz Bornmann, and Robin Haunschild.
Published under a Creative Commons
Attribution 4.0 International (CC BY 4.0)
license.



RESEARCH ARTICLE

Can altmetrics reflect societal impact considerations?: Exploring the potential of altmetrics in the context of a sustainability science research center

Omar Kassab¹, Lutz Bornmann² , and Robin Haunschild³ 

¹Swiss Federal Institute of Technology (ETH Zurich), Professorship for Social Psychology and Research on Higher Education, Andreasstrasse 15, 8050 Zurich, Switzerland

²Administrative Headquarters of the Max Planck Society, Division for Science and Innovation Studies, Hofgartenstrasse 8, 80539 Munich, Germany

³Max Planck Institute for Solid State Research, Heisenbergstrasse 1, 70569 Stuttgart, Germany

Keywords: altmetrics, bibliometrics, MHq indicator, research center, societal impact, sustainability science

ABSTRACT

Societal impact considerations play an increasingly important role in research evaluation. In particular, in the context of publicly funded research, proposal templates commonly include sections to outline strategies for achieving broader impact. Both the assessment of the strategies and the later evaluation of their success are associated with challenges in their own right. Ever since their introduction, altmetrics have been discussed as a remedy for assessing the societal impact of research output. On the basis of data from a research center in Switzerland, this study explores their potential for this purpose. The study is based on the papers (and the corresponding metrics) published by about 200 either accepted or rejected applicants for funding by the Competence Center Environment and Sustainability (CCES). The results of the study seem to indicate that altmetrics are not suitable for reflecting the societal impact of research that was considered: The metrics do not correlate with the ex ante considerations of an expert panel.

1. INTRODUCTION

Many studies dealing with the societal impact of research begin by describing a paradigmatic transformation in research policy that has presumably led to an increased accountability of publicly funded research. Researchers and universities, according to this narrative, would increasingly have to justify their work to the tax-paying public. This almost confrontational portrayal of the relationship could make the reader believe that the public is concerned with a petty cost–benefit calculation to tease out their return on investment. However, this simplified view undermines the potentially genuine interest of societal actors to inform and educate themselves on the basis of scientific facts. Especially in times of rapid technological developments, the interaction between science and society is easier than ever.

The very emergence of social media, for example, has heralded a new age for the public dissemination of scientific knowledge. It therefore comes as no surprise that “altmetrics,” an endeavor to quantitatively represent mentions and interactions on social media platforms such

as Twitter or Facebook, have been proposed as a means to evaluate the societal impact of research ex post (see the literature overview by Bornmann, 2014). Yet despite the consensus over their potential for impact assessment, the jury is still out as to what kind of impact altmetrics scores actually reflect. Addressing this puzzle, Bornmann, Haunschild, and Adams (2019) compared peer assessments of societal impact of research with altmetrics scores for the corresponding publications. Their results reveal that altmetrics seem to measure public “discussions” around research rather than societal impact, further qualifying that the latter may more likely be assessed by experts in a specific field. However, there are also other empirical findings suggesting a contrary conclusion. Wooldridge and King (2019), for example, used the same data set as Bornmann et al. (2019) but other methods, and concluded that “the work presented in this study provides direct evidence, for the first time, of a correlation between expert peer review of the societal impact of research and altmetric data from the publications defining the underpinning research” (p. 281). Against the backdrop of these contradicting results, it is necessary to advance further empirical investigations about the correlation between assessments of societal impact of research and altmetrics scores.

Taking up the question in the context of a research center, the CCES in Switzerland, the study examines the altmetrics scores of journal articles published by researchers either accepted or rejected for funding by CCES. As a research field “defined by the problems it addresses rather than by the disciplines it employs” (Clark, 2007), sustainability science represents a prime case for solution-oriented research of high societal relevance (Yarime, Trencher, et al., 2012; Brandt, Ernst, et al., 2013; Wiek, Talwar, et al., 2014; Kassab, 2019). Thus, whether the research was funded or rejected depended not solely on the assessment of the scientific quality, but initially on whether the prospect of societal impact was explicitly outlined in the proposal or not (CCES, personal communication). We explore in this study whether this latter criterion is reflected in later altmetrics scores: Do papers of researchers funded by CCES receive higher altmetrics scores than papers from rejected researchers? Or in other words, using another data set than Bornmann et al. (2019) and Wooldridge and King (2019), this study targets the question of whether or not altmetrics scores are consistent with ex ante assessments of societal impact considerations.

The remainder of the article is structured as follows: Section 2 introduces the case and describes the hypothesized relationship between societal impact assessments and altmetrics scores. Section 3 then gives an overview of the data and the methods used for the investigation. Section 4 presents the results of the study, and section 5 discusses them to draw conclusions. Finally, section 6 outlines the limitations of the study while giving indications for further research and recommendations.

2. CAN ALTMETRICS REFLECT SOCIETAL IMPACT CONSIDERATIONS?

2.1. Case Description: A Sustainability Science Research Center in Switzerland

The CCES was founded in 2006 for a period of 10 years (until 2016) to foster inter- and transdisciplinarity within and between the six institutions that constitute the ETH Domain, a union of Swiss Federal universities and research institutes. Strategically managed by the ETH Board, the ETH Domain comprises the two Federal Institutes of Technology in Zurich (ETH Zurich) and Lausanne (EPFL), as well as four research institutes: the Paul Scherrer Institute (PSI), the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), the Swiss Federal Laboratories for Materials Science and Technology (Empa), and the Swiss Federal Institute of Aquatic Science and Technology (Eawag).

CCES was established with the mission to “identify the relevant questions and the appropriate answers to foster the sustainable development of a future society while minimizing the impact on the environment” (CCES, 2005). To comprehensively achieve this mission, CCES operated in three areas of activity: research, capacity-building, and public outreach. Goals have been set for each of the three areas, with a total of five goals. In the area of “research,” three goals were defined: (a) foster major inter- and transdisciplinary research advancements in the areas of environment and sustainability, (b) establish the CCES partner institutions as national and international focal points for the areas of environment and sustainability, and (c) achieve a long-term structuring effect and a coherent strategy for the areas of environment and sustainability. In the area of “capacity-building,” the goal was (d) to establish a strong and wide-ranging education program for the areas of environment and sustainability. And lastly, the goal set in the area of “public outreach” was (e) to achieve a visible societal impact with a focus on socioeconomic implementation.

Activities at CCES were clustered in 26 projects along five thematic areas of environment and sustainability science: (a) Climate and Environmental Change, (b) Sustainable Land Use, (c) Food, Environment, and Health, (d) Natural Resources, and (e) Natural Hazards and Risks. Some exemplary projects included OPTIWARES, in which researchers worked on optimizing the use of wood as a renewable energy source, TRAMM, which aimed at developing early warning systems for rapid mass movements in steep terrain, and the ADAPT project, which studied social and environmental constraints for large-scale dams and water resource management (Kassab, Schwarzenbach, & Gotsch, 2018).

2.2. Societal Impact Considerations in the Evaluation Procedure

The few aforementioned synopses demonstrate exemplarily that projects funded by the research center were characterized by a strong practice orientation. This property is based on the notion that sustainability science concentrates on the most pressing challenges facing human society and the development of concrete solutions (Yarime et al., 2012; Kajikawa, Tacoma, & Yamaguchi, 2014; SDSN, 2017). In order to find these solutions, however, it is necessary not only to overcome disciplinary boundaries, through interdisciplinarity, but also transcend the university ecosystem and engage other stakeholders from society, business, and politics, through transdisciplinarity approaches (Pohl, 2010; Lang, Wiek, et al., 2012). In terms of the underlying research mode, sustainability science thus differs considerably from basic research (Clark, 2007; Möbjörk, 2010; Kates, 2011; Miller, 2013).

The special attention given to inter- and transdisciplinarity as well as the objective to develop applied solutions was explicitly reflected in the CCES evaluation procedure. For the purpose of assessing the project proposals, an ad hoc Research Council (RC) was established. Consisting of 17 researchers from the ETH Domain institutions, the RC was responsible for reviewing proposals with respect to their overall suitability for CCES (see goals above). In particular, it was the task of the RC to evaluate the added value of the project for CCES, stressing (a) societal relevance, either as a goal to be achieved during the project duration or with an identified follow-up implementation phase, (b) the importance of the project for long-term sustainability and for a durable structuring effect, and (c) the relevance in the international context, and in particular, the potential for applications in developing countries (CCES, 2006). As this focus suggests, the assessments of the RC were primarily based on the prospect of societal impact, reflecting the three aforementioned dimensions, and did not include an evaluation of the scientific quality. In fact, only if the projects passed the initial assessment were they forwarded to the next stage, which consisted of a classical peer review procedure coordinated by the ETH Zurich Research Commission.

Given the still inconclusive debate about the validity of altmetrics for reflecting the societal impact of research, the question that lies at the heart of this study is whether or not there is a relationship between ex ante assessments of societal impact and altmetrics scores. We approach the answer to this question indirectly: According to the CCES evaluation procedure outlined above, special emphasis was attributed to the prospect of societal impact. Under the premise that research funded through CCES would yield more societal impact than the research of rejected applicants, and assuming that altmetrics scores are capable of reflecting this impact, the hypothesis arises that the *researchers funded by CCES achieve higher impact in terms of altmetrics scores with their research than those who were not funded*. Should the findings of this study corroborate the hypothesis, this would lead to the conclusion that altmetrics are indeed capable of reflecting the ex ante societal impact considerations of the RC. However, should the results not confirm the hypothesis, this does not automatically imply the opposite. Rather, this would raise the question of what else altmetric scores are indicative of. In fact, a refutation of the hypothesis could also be interpreted in a way that the RC did not take sufficient account of societal impact considerations in the assessments (even though this was explicitly demanded), but rather focused on other aspects. In what follows, we describe the data and the methods we use to test the hypothesis empirically.

3. DATA AND METHODS

3.1. Description of Altmetrics

We acknowledge that altmetrics are heterogeneous in many ways and specifically with regard to which aspect of societal impact they actually reflect (if any). We considered six different altmetric sources in this study, including Twitter, Wikipedia, policy-related documents, blogs, Facebook, and news. They differ strongly with regard to the effort and the process preceding the actual mention, content, and substance of the information that is communicated, and also the readership. While a tweet or a Facebook post is shared at the touch of a button, the threshold for Wikipedia entries, blog posts, or mentions in news outlets is much higher. Also, the demographic background of the readership of policy-related documents as opposed to Facebook posts is much more specific. Nevertheless, we chose those types since they have been frequently used and investigated in previous altmetrics studies (see Bornmann et al., 2019), qualifying them as “standard” sources.

Twitter (<https://www.twitter.com>) is a popular microblogging platform. Tweets may refer to the content of scientific publications, but it seems that they do not correlate with traditional citations (Bornmann, 2015). Instead, they may reflect discussion around these publications (Haustein, Peters, et al., 2014), possibly by public users (Haustein, Larivière, et al., 2014; Yu, 2017), but this is not entirely clear, as outlined by Sugimoto, Work, et al. (2016). The results by Andersen and Haustein (2015) suggest that tweets reflect the attractiveness of papers for a broader audience. However, contradictory results are also available: “A multi-year campaign has sought to convince us that counting the number of tweets about papers has value. Yet, reading tweets about dental journal articles suggested the opposite. This analysis found: obsessive single issue tweeting, duplicate tweeting from many accounts presumably under centralized professional management, bots, and much presumably human tweeting duplicative, almost entirely mechanical and devoid of original thought” (Robinson-Garcia, Costas, et al., 2017). In the study at hand, the number of tweets (and retweets) including references to scientific papers in our data set is counted.

Wikipedia (<https://www.wikipedia.org>) is a free encyclopedia platform which includes editable content (Mas-Bleda & Thelwall, 2016). Although contributors to this platform include scholarly references, most of them do not refer to research papers (Priem, 2014). If scientific papers are cited, open access (OA) papers seem to be preferred (Teplitskiy, Lu, & Duede, 2017; Dehdarirad, Didegah, & Sotudeh, 2018). Guglielmi (2018) reports on Wikipedia's most frequently mentioned papers. However, this list does not correspond with lists based on traditional citations: Study results suggest that Wikipedia mentions do not correlate with citations (Samoilenko & Yasseri, 2014). A Wikipedia case study with papers on wind power showed that less than 1% of relevant papers have been cited on Wikipedia, "implying that the direct societal impact through the Wikipedia is extremely small for Wind Power research" (Serrano-López, Ingwersen, & Sanz-Casado, 2017, p. 1471). Kousha and Thelwall (2017) found that only 5% of papers had any citation from Wikipedia—based on a significantly larger sample of papers than considered by Serrano-López et al. (2017). In this study, the number of Wikipedia articles with reference to papers in our data set is counted.

Policy-related documents are an important source of altmetrics, since one is interested in the impact of science on the policy realm (OPENing UP, 2016; Vilkins & Grant, 2017). Mentions in these documents are searched using text mining databases of, for instance, the World Health Organization or European Food Safety Authority (Bornmann, Haunschild, & Marx, 2016; Haunschild & Bornmann, 2017). Haunschild and Bornmann (2017) reported that the company Altmetric tracks more than 100 policy sources (in 2015). Tattersall and Carroll (2018) analyzed nearly 100 papers published by authors from the University of Sheffield: The "research topics with the greatest policy impact are medicine, dentistry, and health, followed by social science and pure science." Papers published OA seem to have an advantage to be cited in policy-related documents (Vilkins & Grant, 2017). However, the impact of papers (OA or not) on these documents is usually very low, as the results of Haunschild and Bornmann (2017) reveal: "Less than 0.5% of the papers published in different subject categories are mentioned at least once in policy-related documents" (p. 1209). The study of Bornmann et al. (2016) shows that "only 1.2 % ($n = 2,341$) have at least one policy mention" (p. 1477). The authors analyzed a large set of 191,276 publications from the field of climate change, which is policy relevant. In this study, the number of policy-related documents with references to papers in our data set is counted.

Blogs are written about scientific papers, including formal or informal citations of papers (Shema, Bar-Ilan, & Thelwall, 2014a). These citations can be counted—with the limitation that informal citations lead to uncertainty (Priem & Hemminger, 2010; Luzón, 2013; Shema, Bar-Ilan, & Thelwall, 2014b). Since blogs allow extended informal discussions about research, they are an interesting altmetrics source (Fausto, Machado, et al., 2012; Shema, Bar-Ilan, & Thelwall, 2012a). Blogging may be a bridge between the general public and the research area (Bonetta, 2007; Bar-Ilan, Shema, & Thelwall, 2014), whereby bloggers seem to have preferences for papers from high-impact journals and research in the life and behavioral sciences (Shema, Bar-Ilan, & Thelwall, 2012b). However, a study revealed that bridging public and research "was one of the less popular motivations for academics to blog" (Mewburn & Thomson, 2013, p. 1113). The literature overview published by Sugimoto et al. (2016) shows that the coverage of papers in blog mentions is low, as is the correlation between blog mentions and traditional citations. In this study, the number of blog posts with references to the papers in our data set is counted.

Facebook is a popular social networking and social media platform (Bik & Goldstein, 2013). Since users share papers among themselves, mentions of papers in posts or Facebook likes can be counted. Ringelhan, Wollersheim, and Welpé (2015) investigated

whether Facebook “likes” are an indicator of scientific impact. Their results show “an interdisciplinary difference in the predictive value of Facebook likes, according to which Facebook likes only predict citations in the psychological area but not in the nonpsychological area of business or in the field of life sciences.” In this study, the number of Facebook posts with references to scientific papers in our data set are counted (note that we did not include likes).

News attention relates to scientific papers mentioned in news reports (via direct links or unique identifiers in, for example, the *New York Times*). On the basis of these paper mentions, public attention can be counted. The overview of altmetrics studies published by Sugimoto et al. (2016) reveals that the correlation between mentions of papers in news reports and traditional citations is between low and medium. According to *Altmetric.com*¹, more than 2,000 different news sources are analyzed for news mentions. In this study, the number of news articles with mentions of scientific papers in our data set is counted.

3.2. Data Set Used

We used the Web of Science (WoS, Clarivate Analytics) custom data of our in-house database and the database from the Competence Centre for Bibliometrics (CCB: <http://www.bibliometrie.info>). Both are derived from the Science Citation Index Expanded (SCI-E), Social Sciences Citation Index (SSCI), and Arts and Humanities Citation Index (AHCI) produced by Clarivate Analytics (Philadelphia, USA). All publications published between 2011 and 2015 with a DOI were exported with the following information: DOI, WoS UT (unique accession number from WoS), WoS subject categories, publication year, citation counts with a three-year citation window starting after the publication year, and Hazen percentiles. Percentiles are field- and time-normalized impact scores that are between 0 (low citation impact) and 100 (high citation impact) (Bornmann, Leydesdorff, & Mutz, 2013). Raw citation data were taken from the database maintained by the CCB. Both databases (our in-house database and the database maintained by the CCB) were last updated at the end of April 2019. We kept only those publications that fulfilled the following two criteria: (a) the publication belongs to a field (overlapping WoS category; Rons, 2012, 2014) to which at least one research center publication belongs; (b) a requirement of at least 10 publications per field and publication year combination has been set.

Altmetrics data were sourced from a locally maintained database using data shared with us by Altmetric (<https://www.altmetric.com>) and dumped on October 8, 2019. For research projects, the company shares the data for free. The data include altmetric counts from sources such as social networking, blogging, microblogging, wikis, and policy-relevant usage. We appended a mention count to each DOI using the following altmetrics sources: Twitter, Facebook, blogs, news, policy documents, and Wikipedia (see above). One DOI not known to the altmetrics database was recorded as “not mentioned.” Altmetrics data and information about their unit status (applied for research center funding which was accepted or not) were appended to the publications via their DOI. Figure 1 provides a schematic overview of how the respective units were constructed: unit 0 contains all WoS papers that do not belong to units 1 or 2. Unit 1 contains the publications of 28 participants who had submitted project proposals for CCES but were not funded. Unit 2, in turn, contains the publications of 170 participants that were affiliated with CCES as principal investigators and project partners. Unit 2 is

¹ <https://www.altmetric.com/about-our-data/our-sources/>.

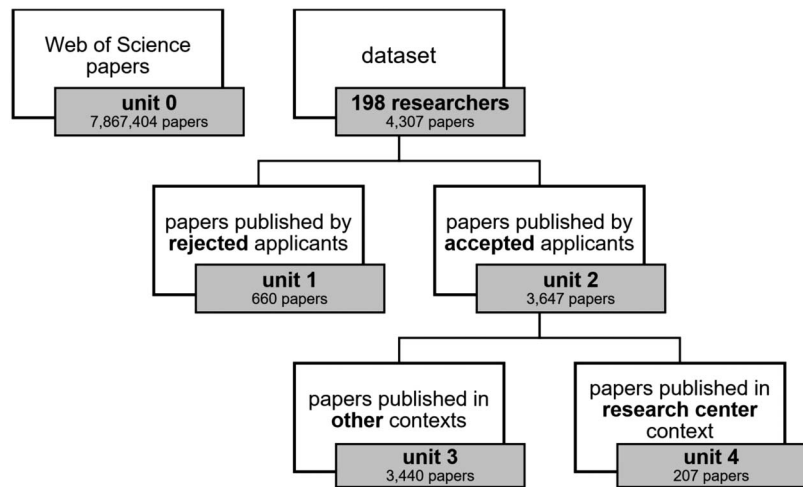


Figure 1. Schematic overview of the units and number of papers per unit.

further subdivided into units 3 and 4. Unit 4 contains the papers that were published in the research center context, while unit 3 contains papers that accepted applicants published beyond their project at the research center. The numbers of mentioned and not mentioned publications in the different altmetrics sources broken down by unit status and publication year are shown in Table 1.

We acknowledge that the subdivision into units and the comparison between units is a simplification of reality, especially with regard to the hypothesis to be tested. While the CCES evaluation procedure took place on the project level at a specific moment in time, the units here are constructed on the level of the entire publication output of researchers that were funded or not funded by CCES. Furthermore, we focus in this study on scientific publications as the main research output of the research center. While it would have been beneficial to consider other outputs as well, such as those emanating from public outreach activities (Kassab, 2019), we are constrained by the fact that altmetrics data are only available for outputs that have a unique and persistent identifier, such as a DOI. However, besides altmetrics data, we also considered citation data (a) to compare the results with those based on altmetrics data and (b) to investigate whether societal impact assessments correspond with traditional impact scores.

3.3. Mantel-Haenszel Quotient (MHq)²

In this study, we compare the impact of papers published by various units (e.g., papers published by rejected or accepted applicants; see Figure 1). Since altmetrics data concern field-specific differences (like citation data), field-normalized indicators should be used instead of raw data for group comparisons. However, it is a critical drawback of altmetrics data that they are inflated by zeros: In the current study, 5,586,077 papers (71%) have no impact in any altmetrics source. For zero-inflated data, it is not possible to use methods for field normalization that are usually applied in bibliometrics (methods based on mean citations or citation percentiles; Bornmann et al., 2013). Since Bornmann and Haunschild (2018) and

² The explanation of the MHq indicator has been mainly adopted from Bornmann and Haunschild (2018) and Haunschild and Bornmann (2018).

Table 1. Number of mentioned and not mentioned (and cited and not cited) papers, respectively, broken down by data source, publication year, and funded or not-funded groups

Unit	Publication year	Twitter		Facebook		Blogs		News		Policy documents		Wikipedia		<QSS> s	
		Mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Mentioned	Not mentioned	Cited	Not cited
0	2011	139,766	1,233,952	34,842	1,338,876	35,790	1,337,928	20,922	1,352,796	29,028	1,344,690	28,036	1,345,682	1,057,600	316,118
	2012	312,073	1,181,879	77,809	1,416,143	43,989	1,449,963	29,490	1,464,462	28,391	1,465,561	28,024	1,465,928	1,149,376	344,576
	2013	398,950	1,194,767	115,234	1,478,483	52,390	1,541,327	47,867	1,545,850	28,422	1,565,295	27,288	1,566,429	1,245,496	348,221
	2014	510,575	1,144,363	105,419	1,549,519	56,832	1,598,106	59,324	1,595,614	24,927	1,630,011	25,303	1,629,635	1,305,474	349,464
	2015	627,508	1,123,571	17,5196	1,575,883	60,213	1,690,866	76,049	1,675,030	20,868	1,730,211	24,118	1,726,961	1,394,085	356,994
1	2011	27	99	8	118	15	111	4	122	17	109	5	121	120	6
	2012	58	94	25	127	23	129	9	143	21	131	5	147	139	13
	2013	59	113	21	151	20	152	18	154	13	159	7	165	163	9
	2014	61	87	8	140	14	134	19	129	16	132	5	143	146	2
	2015	35	27	17	45	7	55	8	54	5	57	1	61	61	1
2	2011	75	690	15	750	39	726	15	750	46	719	23	742	734	31
	2012	179	650	36	793	52	777	16	813	42	787	21	808	788	41
	2013	285	611	74	822	75	821	53	843	40	856	27	869	862	34
	2014	301	511	57	755	59	753	51	761	25	787	12	800	782	30
	2015	165	180	51	294	24	321	30	315	8	337	7	338	336	9
3	2011	72	627	14	685	34	665	14	685	39	660	22	677	668	31
	2012	173	618	36	755	47	744	14	777	39	752	21	770	750	41
	2013	271	560	70	761	68	763	47	784	31	800	26	805	797	34
	2014	294	489	56	727	58	725	51	732	23	760	12	771	755	28
	2015	164	172	51	285	24	312	30	306	8	328	7	329	328	8
4	2011	3	63	1	65	5	61	1	65	7	59	1	65	66	0
	2012	6	32	0	38	5	33	2	36	3	35	0	38	38	0
	2013	14	51	4	61	7	58	6	59	9	56	1	64	65	0
	2014	7	22	1	28	1	28	0	29	2	27	0	29	27	2
	2015	1	8	0	9	0	9	0	9	0	9	0	9	8	1

Note: WoS papers (unit 0; neither accepted, nor rejected for funding), papers published by rejected applicants (unit 1), and papers published by accepted applicants (unit 2). The papers from accepted applicants are further divided into papers from funded projects (unit 4) and papers published in other contexts (unit 3).

Haunschild and Bornmann (2018) proposed the MHq indicator, which is specifically designed for dealing with zero-inflated data in field normalization, we used the indicator in the current study.

For pooling data from multiple 2×2 cross tables based on such subgroups (which are part of the larger population, including all papers in the considered time period), MH analysis is a popular method (Mantel & Haenszel, 1959; Hollander & Wolfe, 1999; Sheskin, 2007). According to Fleiss, Levin, and Paik (2003), the method “permits one to estimate the assumed common odds ratio and to test whether the overall degree of association is significant. Curiously, it is not the odds ratio itself but another measure of association that directly underlies the test for overall association ... The fact that the methods use simple, closed-form formulas has much to recommend it” (p. 250). The results by Radhakrishna (1965) demonstrate that the MH approach seems to be valid.

The MH analysis results in a summary odds ratio for multiple 2×2 cross tables, which Bornmann and Haunschild (2018) and Haunschild and Bornmann (2018) name *MHq*. For the comparison of the papers published by the applicants with reference sets in view of impact, the 2×2 cross tables (which are pooled) consist of the number of papers mentioned and not mentioned in subject category and publication year combinations f . In the 2×2 subject-specific cross table (see Table 2), the cells a_f , b_f , c_f , and d_f are defined as follows:

- a_f is the number of mentioned papers published by unit g (e.g., rejected applicants) in subject category and publication year f ,
- b_f is the number of not mentioned papers published by unit g in subject category and publication year f ,
- c_f is the number of mentioned papers in subject category and publication year f , and
- d_f is the number of not mentioned papers published in subject category and publication year f . Note that the papers of group g are also part of the papers in the world.

The following dummy variables are needed for the MH analysis:

$$R_f = \frac{a_f d_f}{n_f} \quad \text{and} \quad R = \sum_{f=1}^F R_f, \quad (1)$$

$$S_f = \frac{b_f c_f}{n_f} \quad \text{and} \quad S = \sum_{f=1}^F S_f, \quad (2)$$

$$P_f = \frac{a_f + d_f}{n_f} \quad \text{and} \quad Q_f = 1 - P_f \quad (3)$$

where $n_f = a_f + b_f + c_f + d_f$.

MHq is simply

$$\text{MHq} = \frac{R}{S} \quad (4)$$

The CIs for MHq are calculated following Fleiss et al. (2003). The variance of $\ln \text{MHq}$ is estimated by

$$\widehat{\text{Var}}(\ln \text{MHq}) = \frac{1}{2} \left\{ \frac{\sum_{f=1}^F P_f R_f}{R^2} + \frac{\sum_{f=1}^F (P_f S_f + Q_f R_f)}{RS} + \frac{\sum_{f=1}^F Q_f S_f}{S^2} \right\} \quad (5)$$

Table 2. 2×2 subject-specific cross table

	Number of mentioned papers	Number of not mentioned papers
Group g	a_f	b_f
World	c_f	d_f

The CI for the MHq can be constructed with

$$MHq_L = \exp \left[\ln(MHq) - 1.96 \sqrt{\widehat{Var}[\ln(MHq)]} \right] \quad (6)$$

$$MHq_U = \exp \left[\ln(MHq) + 1.96 \sqrt{\widehat{Var}[\ln(MHq)]} \right] \quad (7)$$

We used the data in Table 3 to produce a small world example for explaining the MHq: The world consists of papers in four subject categories. The papers of two units (publication set A and B) determine the world. For each unit, the numbers of mentioned and not mentioned papers as well as the corresponding proportion of mentioned papers are listed. For example, the unit named as publication set B has published 26 mentioned and seven not mentioned papers in subject category 1. The proportion of the papers mentioned is 0.27. It is an advantage of the MHq that the world average has a value of 1: This value indicates that there is no difference between the chances of a focal publication set and the reference sets (i.e., the world) of being mentioned (e.g., on Wikipedia). An MHq value less than 1.0 indicates lower chances for the publications in the set of being mentioned compared with the reference sets. The MHq values in Table 3 can be interpreted as follows: The chances of the papers in publication set A of being mentioned are 0.81 times as large as the world's papers' chances. The chances of the papers in set B of being mentioned are 1.3 times greater than the world's papers' chances. It is an advantage of the MHq that the result can be expressed as a percentage, which is relative to the world average. Expressed as percentages, therefore, the difference between publication set B and the world is

$$100 * (1.3 - 1.0) = 30\% \quad (8)$$

Thus, the publications in set B have a 30% higher chance of being mentioned than the world's publications. We also added CIs to the MHqs in Table 3. Since the CIs of both publication sets (A and B) overlap substantially among themselves and with 1.0 (the world MHq), they do not differ statistically significantly from one another and the world average.

4. RESULTS

Figure 2 displays the MHq values (based on six altmetrics sources) for all WoS papers in the given years (unit 0: red points; neither accepted nor rejected); papers published by rejected applicants (unit 1: green squares); and papers published by accepted applicants (unit 2: blue diamonds). The papers from accepted applicants are further differentiated into papers written in the context of projects funded by the research center (unit 4: orange diamonds) and papers published in other contexts (unit 3: yellow diamonds). For all MHq values, CIs are indicated. Since the paper numbers from funded projects for some publication years are too low, they could not be presented in the figure.

Table 3. Small world example for the Mantel–Haenszel quotient (MHq)

World (reference sets)	Paper is mentioned	Paper is not mentioned	Number of papers	MHq
Subject category 1	44	20	64	
Subject category 2	30	16	46	
Subject category 3	16	12	28	
Subject category 4	0	20	20	
Total				1.00 [0.61, 1.64]
Publication set A				
Subject category 1	18	13	31	
Subject category 2	15	9	24	
Subject category 3	13	9	22	
Subject category 4	0	10	10	
Total				0.81 [0.46, 1.44]
Publication set B				
Subject category 1	26	7	33	
Subject category 2	15	7	22	
Subject category 3	3	3	6	
Subject category 4	0	10	10	
Total				1.30 [0.66, 2.53]

The results as summarized in Figure 2 do not support the hypothesis that funded researchers achieve higher altmetrics scores with their research than those who were not funded by the research center. For example, the MHq values based on Twitter data for the papers published by rejected applicants (green squares) are consistently higher than the papers published by accepted applicants (blue diamonds). The differences between both groups are statistically significant in 2011 and 2012, but not in 2013 to 2015 (here, the CIs mostly overlap). Quite strikingly, the figure also reveals that papers published by accepted applicants in the context of the funded research center projects (orange diamonds) even receive lower Twitter scores than the papers they published outside of the research center project (yellow diamonds). The results for the other altmetric scores mainly concur with the Twitter results. Only the findings for the policy-related documents show a different picture: Research-center-based papers published between 2012 and 2014 (orange diamonds) received higher altmetric scores than papers by the same researchers that do not emanate from research center projects (yellow diamonds). However, the results are not statistically significant and are not confirmed by the results for 2011 (results for 2015 are not available).

We further analyzed whether the ex ante societal impact considerations are reflected in citation scores. The results are shown in Figure 3. The figure reveals that the results are more

or less in agreement with the altmetrics results (with papers published by rejected applicants performing similarly to or better than those of funded applicants). If we inspect the aggregated MHq results based on the papers from all years, papers published by accepted applicants (MHq = 3.31) have a higher citation impact than papers published by rejected applicants (MHq = 2.87). Since the CIs of both groups overlap, however, the results are not statistically significant. We obtained similar results (missing substantial differences between the groups), when we compared median citations (accepted applicants = 9, rejected applicants = 9) and percentile citation scores (accepted applicants = 73, rejected applicants = 70) of both groups.

5. DISCUSSION AND CONCLUSION

Universities and researchers are increasingly under pressure to disclose how their research contributes to the welfare of society to garner political support and funding (Puschmann, 2014; Thune, Reymert, et al., 2016; Bornmann & Haunschild, 2017). In light of this development, assessing the societal impact of research is a critically debated issue among evaluation scholars and research policy experts. Because of their widespread use, social media have been at the heart of methodological discussions over the past years, including both their potential (e.g., speed, broadness) and their shortcomings (e.g., data quality, zero-inflated data). However, the critical question of whether social media (or altmetrics for that matter) are able to reflect societal impact is so far not answered due to conflicting empirical evidence. Against this background, the aim of this study was to contribute to solving this puzzle. For this purpose, the present paper compared ex ante assessments on the societal relevance of research with altmetrics scores that the respective research received in the years after. A research center from the field of sustainability science (CCES) and the societal impact assessments made by the members of an ad hoc RC served as the case study for this investigation.

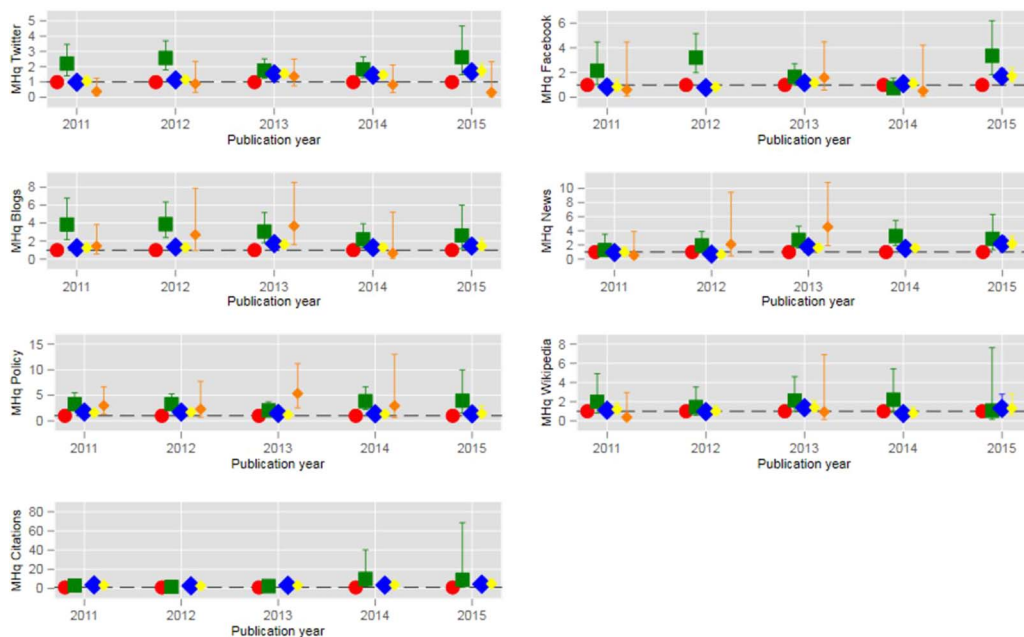


Figure 2. MHq values based on six altmetrics sources for all WoS papers (unit 0: red points; neither accepted, nor rejected); papers published by rejected applicants (unit 1: green squares); and papers published by accepted applicants (unit 2: blue diamonds). The papers from accepted applicants are further divided into papers from funded projects (unit 4: orange diamonds) and papers published in other contexts (unit 3: yellow diamonds). For some years, the values of unit 4 are missing because the numbers of mentioned papers are too low.

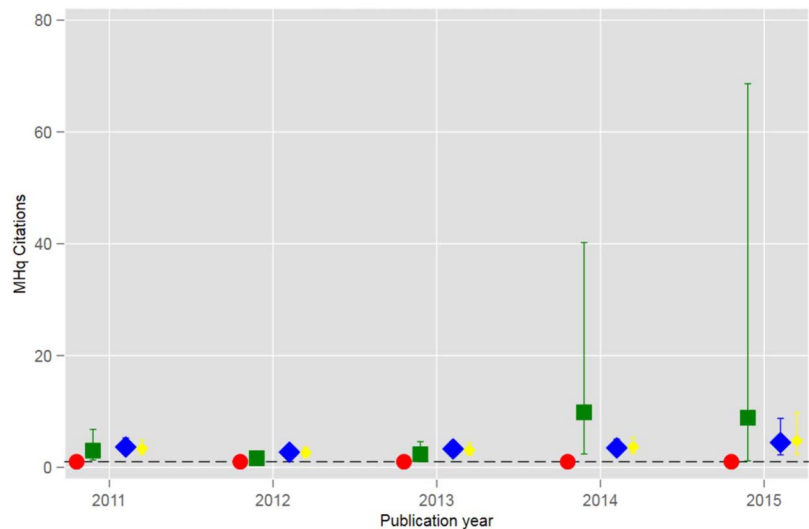


Figure 3. MHq values based on citation counts for all WoS papers (unit 0: red points; neither accepted nor rejected for funding), papers published by rejected applicants (unit 1: green squares), papers published by accepted applicants (unit 2: blue diamonds), and papers published in other contexts from accepted applicants (unit 3: yellow diamonds). Papers published from funded projects (unit 4) are not shown, because the numbers of uncited papers are too low.

In conclusion, the proposed hypothesis that researchers funded by CCES achieve higher impact in terms of altmetrics scores with their research than those who were not funded could not be confirmed based on the results. We found no correlation between the RC's assessment and the corresponding altmetric scores. With a few exceptions, this finding seems to be confirmed in the case of all six types of altmetrics. For comparison with altmetrics, we investigated the relationship with citation scores as well. The results are similar to those based on altmetrics: The correlation is not in the expected direction.

Our results might be interpreted in such a way that altmetrics are not entirely suitable for reflecting the societal impact of research. However, since we investigated only one case-specific evaluation procedure and the results are not homogeneous throughout the different types of altmetric scores, this conclusion cannot be drawn with certainty. We conclude therefore that more research is needed to better understand what altmetrics are reflecting, particularly in light of their heterogeneity. Further research should clarify whether altmetric scores rather capture "unknown attention or unstructured noise produced by published research" (Moed, 2017; Bornmann et al., 2019), some sort of "public discussion" (Haunschild, Leydesdorff, et al., 2019) or anything else altogether.

Our results, at the same time, could be interpreted with a critical view of the RC's assessments. Did the members of the RC select the "right" projects in the first place or how should the missing correlation between the ex ante assessments and the received citations be interpreted?

Another question is whether the members of the RC were qualified to judge the societal impact of proposed research. In most cases, expert panels are composed only of researchers rather than of representatives of other sectors of society, which was also true for in the case of CCES. This circumstance may have led to the fact that the potential societal impact could not be accurately judged or that the aspect of societal impact was not given enough importance in

the evaluation procedure. Overall, we note that our findings can take the discussion forward, but also that they should be interpreted with caution.

6. LIMITATIONS, FURTHER RESEARCH, AND RECOMMENDATIONS

The study revealed that the *ex ante* assessments considering societal impact of research and the altmetric scores of the same research do not correlate. We could conclude the debate at this point and throw altmetrics overboard as a potential measure of societal impact. But, of course, this study has several limitations that need to be discussed.

One key aspect is related to the fact that altmetrics are still in their infancy in many ways. For example, are altmetrics really a good proxy for societal impact? Are social media mentions in themselves societal impact? Does a mention or interaction on social media automatically imply that there has been a cognitive engagement with the content of the research, and that societal impact has occurred? Or is it perhaps a buzzword-laden title, zeitgeist, or fame-related reason why some research output scores higher on altmetrics than others (Hall, 2014)? And what can we say about all the research that does not have any mentions on Twitter or Wikipedia? It would be highly questionable to conclude that no societal impact has been achieved in all these null observation cases. Furthermore, our study does not differentiate between “self-mentions” or “in-house users” (the own department or the university’s communications team) and mentions and interactions by other (more independent) individuals and entities. Despite being somewhat complex, further research should account for these aspects, as well as for the actual content of the tweets or Facebook posts. This latter strategy could allow for a better understanding of the intentions and meanings of social-media-based interactions with research. By looking at the content or the timing of the mentions in more detail, we could possibly identify different strategies in using social media, which could help us formulate new hypotheses.

The results of this work have again shown that the true value added of altmetrics is not yet entirely clear, but rather ranges on a scale between societal impact and unstructured noise. This fundamental problem concerns all six types of altmetrics that have been considered in this study (to a greater or lesser extent). With regard to the inability of tweets to measure the societal impact of research, the results of the present study are consistent with those of Haustein et al. (2014) and Andersen and Haustein (2015). From our point of view, off-the-cuff retweets in particular are simply too inflationary to imply a serious engagement with the content of the work. Mentions on Wikipedia also do not seem to reflect the societal relevance of research (Kousha & Thelwall, 2017). Further, it does not yet seem to be common practice to incorporate scientific research into policy and policy-related documents, either in the field of climate change, as Bornmann et al. (2016) found, or in the likewise societally relevant field of sustainability science, as the present study showed. This finding also underlines that the dialog and knowledge transfer between science and policy is far from established (Hessels, Van Lente, & Smits, 2009). At the same time, it must be given fair consideration that, as with classical citations, it can take up to several years for the results of scientific studies to become relevant and cited in policy documents. The time window of the present study was simply too small. Finally, the valid assessment of societal impacts by means of blogging, Facebook, or in news outlets largely suffers from the fact that there is a bias toward publications from renowned journals (Shema et al., 2012a, 2012b) or specific fields of interest (Ringelhan et al., 2015). Although our findings seem to lend additional support in favor of the argument that altmetrics are not capable of reflecting the societal impact of research, much more research will need to be done before we can actually have a clear picture of what altmetrics are capable of.

Another limitation of this study is related to the evaluation process itself, and less to the shortcomings of altmetrics. Even though the prospect of societal impact was a key criterion for the RC, the assessments were not based on standardized rating scales along individual criteria, but rather on a holistic rating, we can only assume that societal impact considerations played a role in the evaluation process rather than having clearly traceable evidence for the specific weighting that this critical aspect ultimately had. One remedy for future evaluations could therefore be to assess societal impact as a single dimension using a standardized scale.

A related limitation of the study is associated with the heterogeneity of the societal impact. Societal impact can manifest itself in different ways, such as in the form of policy impact, environmental impact, health impact, or educational impact. Due to the holistic rating in the assessment process, it is not clear what kind of societal impact was the focus of the experts' attention. This heterogeneity is also an issue for altmetrics. A Twitter mention compared to a mention in a policy document, for example, has not only been made a different way but probably has a different kind of impact as well.

With regard to the societal impact of research, this study focused exclusively on the published journal papers and the corresponding altmetrics scores they received. It could certainly bring added value if other outputs were taken into account as well, such as outputs that researchers produce within the framework of public outreach activities. Specifically designed to catalyze the societal impact of research, for example, stakeholder publications or teaching material and their respective altmetrics scores could much more accurately reflect the societal impact of research. However, in order for these alternative types of outputs to receive an altmetrics score, they would have to be assigned a unique and persistent identifier, such as a DOI in the future (see <https://www.doi.org/>).

ACKNOWLEDGMENTS

The bibliometric data used in this study are from two sources: (a) the online version of the Web of Science database provided by Clarivate Analytics (Philadelphia, Pennsylvania, USA) and (b) the bibliometric in-house databases of the Max Planck Society (MPG) and the Competence Centre for Bibliometrics (CCB: <http://www.bibliometrie.info/>). The MPG's database is developed and maintained in cooperation with the Max Planck Digital Library (MPDL, Munich), and the CCB's database is developed and maintained by the cooperation of various German research organizations. Both databases are derived from the Science Citation Index Expanded (SCI-E), Social Sciences Citation Index (SSCI), and the Arts and Humanities Citation Index (AHCI) prepared by Clarivate Analytics, formerly the IP & Science business of Thomson Reuters (Philadelphia, Pennsylvania, USA). The authors thank the management of the CCES for their kind support in conducting this study. The altmetrics data were retrieved from our locally maintained database with data shared with us and dumped by the company [altmetric.com](https://www.altmetric.com) on October 8, 2019.

AUTHOR CONTRIBUTIONS

Omar Kassab: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, writing—original draft, writing—review & editing. Lutz Bornmann: Conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, validation, visualization, writing—original draft, writing—review & editing. Robin Haunschild: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, writing—original draft, writing—review & editing.

COMPETING INTERESTS

The authors have no competing interests. Between 2013 and 2015, OK worked as an executive assistant to the CCES management. Afterward he joined the Professorship for Social Psychology and Research on Higher Education at ETH Zurich, where he conducted this study in collaboration with coauthors LB and RH.

FUNDING INFORMATION

No funding has been received for this research.

DATA AVAILABILITY

For the data sources, see Acknowledgments section. The authors do not have permission to share the data. Information specific to the research center can be drawn from the performance reporting of CCES, available at <https://cces.ethz.ch/about/downloads.html>.

REFERENCES

- Andersen, J. P., & Haustein, S. (2015). Influence of study type on twitter activity for medical research papers. In A. A. Salah, Y. Tonta, A. A. Salah, C. Sugimoto, & U. Al (Eds.), *The 15th Conference of the International Society for Scientometrics and Informetrics*, pp. 26–36. Istanbul, Turkey: ISSI, Bogazici University Printhouse.
- Bar-Ilan, J., Shema, H., & Thelwall, M. (2014). Bibliographic references in Web 2.0. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond Bibliometrics: Harnessing Multi-Dimensional Indicators of Performance*, pp. 307–325. Cambridge, MA: MIT Press.
- Bik, H. M., & Goldstein, M. C. (2013). An introduction to social media for scientists. *PLOS Biology*, 11(4), e1001535. <https://doi.org/10.1371/journal.pbio.1001535>
- Bonetta, L. (2007). Scientists enter the blogosphere. *Cell*, 129(3), 443–445.
- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895–903. <https://doi.org/10.1016/j.joi.2014.09.005>
- Bornmann, L. (2015). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics*, 103(3), 1123–1144.
- Bornmann, L., & Haunschild, R. (2017). Does evaluative scientometrics lose its main focus on scientific quality by the new orientation towards societal impact? *Scientometrics*, 110(2), 937–943.
- Bornmann, L., & Haunschild, R. (2018). Normalization of zero-inflated data: An empirical analysis of a new indicator family and its use with altmetrics data. *Journal of Informetrics*, 12(3), 998–1011. <https://doi.org/10.1016/j.joi.2018.01.010>
- Bornmann, L., Haunschild, R., & Adams, J. (2019). Do altmetrics assess societal impact in a comparable way to case studies? An empirical test of the convergent validity of altmetrics based on data from the UK research excellence framework. *Journal of Informetrics*, 13(1), 325–340. <https://doi.org/10.1016/j.joi.2019.01.008>
- Bornmann, L., Haunschild, R., & Marx, W. (2016). Policy documents as sources for measuring societal impact: How often is climate change research mentioned in policy-related documents? *Scientometrics*, 109(3), 1477–1495. <https://doi.org/10.1007/s11192-016-2115-y>
- Bornmann, L., Leydesdorff, L., & Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits. *Journal of Informetrics*, 7(1), 158–165. <https://doi.org/10.1016/j.joi.2012.10.001>
- Brandt, P., Ernst, A., Gralla, F., Luederitz, C., Lang, D. J., ... von Hehrden, H. (2013). A review of transdisciplinary research in sustainability science. *Ecological Economics*, 92, 1–15.
- CCES. (2005). *Competence Center Environmental and Sustainability (CCES) of the ETH Domain: Business Plan*. Retrieved from: https://ethz.ch/content/dam/ethz/special-interest/dual/cces-dam/documents/CCES_BP_2005_11_22_final.pdf
- Clark, W. C. (2007). Sustainability science: A room of its own. *Proceedings of the National Academy of Sciences*, 104(6), 1737–1738.
- Dehdarirad, T., Didegah, F., & Sotudeh, H. (2018). Which type of research is cited more often in Wikipedia? A case study of PubMed research. In P. Wouters (Ed.), *Proceedings of the Science and Technology Indicators Conference 2018 Leiden "Science, Technology and Innovation Indicators in Transition."* Leiden, Netherlands: University of Leiden.
- Fausto, S., Machado, F. A., Bento, L. F. J., Iamarino, A., Nahas, T. R., & Munger, D. S. (2012). Research blogging: Indexing and registering the change in science 2.0. *PLOS ONE*, 7(12), e50109. <https://doi.org/10.1371/journal.pone.0050109>
- Fleiss, J., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions*, 3rd Ed. Hoboken, NJ: Wiley.
- Guglielmi, G. (2018). Wikipedia's top-cited scholarly articles – revealed. *Nature*, 557(7705), 291–292.
- Hall, N. (2014). The Kardashian index: A measure of discrepant social media profile for scientists. *Genome Biology*, 15(7), 424. <https://doi.org/10.1186/s13059-014-0424-0>
- Haunschild, R., & Bornmann, L. (2017). How many scientific papers are mentioned in policy-related documents? An empirical investigation using Web of Science and Altmetric data. *Scientometrics*, 110(3), 1209–1216. <https://doi.org/10.1007/s11192-016-2237-2>
- Haunschild, R., & Bornmann, L. (2018). Field- and time-normalization of data with many zeros: An empirical analysis using citation and Twitter data. *Scientometrics*, 116(2), 997–1012. <https://doi.org/10.1007/s11192-018-2771-1>

- Haunschild, R., Leydesdorff, L., Bornmann, L., Hellsten, I., & Marx, W. (2019). Does the public discuss other topics on climate change than researchers? A comparison of explorative networks based on author keywords and hashtags. *Journal of Informetrics*, 13(2), 695–707.
- Haustein, S., Larivière, V., Thelwall, M., Amyot, D., & Peters, I. (2014). Tweets vs. Mendeley readers: How do these two social media metrics differ? *IT – Information Technology* 2014, 56(5), 207–215.
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, 101(2), 1145–1163. <https://doi.org/10.1007/s11192-013-1221-3>
- Hessels, L. K., Van Lente, H., & Smits, R. (2009). In search of relevance: The changing contract between science and society. *Science and Public Policy*, 36(5), 387–401.
- Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods*. New York, NY: Wiley.
- Kajikawa, Y., Tacoa, F., & Yamaguchi, K. (2014). Sustainability science: The changing landscape of sustainability research. *Sustainability Science*, 9(4), 431–438.
- Kassab, O. (2019). Does public outreach impede research performance? Exploring the “researcher’s dilemma” in a sustainability research center. *Science and Public Policy*, 46(5), 710–720. <https://doi.org/10.1093/scipol/scz024>
- Kassab, O., Schwarzenbach, R. P., & Gotsch, N. (2018). Assessing ten years of inter- and transdisciplinary research, education, and outreach: The Competence Center Environment and Sustainability (CCES) of the ETH Domain. *GAIA – Ecological Perspectives for Science and Society*, 27(2), 226–234.
- Kates, R. W. (2011). What kind of a science is sustainability science? *Proceedings of the National Academy of Sciences*, 108(49), 19449–19450.
- Kousha, K., & Thelwall, M. (2017). Are Wikipedia citations important evidence of the impact of scholarly articles and books? *Journal of the Association for Information Science and Technology*, 68(3), 762–779. <https://doi.org/10.1002/asi.23694>
- Lang, D. J., Wiek, A., Bergmann, M., Stauffacher, M., Martens, P., ... Thomas, C. J. (2012). Transdisciplinary research in sustainability science: Practice, principles, and challenges. *Sustainability Science*, 7(1), 25–43.
- Luzón, M. J. (2013). Public communication of science in blogs: Recontextualizing scientific discourse for a diversified audience. *Written Communication*. <https://doi.org/10.1177/0741088313493610>
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719–748.
- Mas-Bleda, A., & Thelwall, M. (2016). Can alternative indicators overcome language biases in citation counts? A comparison of Spanish and UK research. *Scientometrics*, 109(3), 2007–2030. <https://doi.org/10.1007/s11192-016-2118-8>
- Mewburn, I., & Thomson, P. (2013). Why do academics blog? An analysis of audiences, purposes and challenges. *Studies in Higher Education*, 38(8), 1105–1119. <https://doi.org/10.1080/03075079.2013.835624>
- Miller, T. R. (2013). Constructing sustainability science: Emerging perspectives and research trajectories. *Sustainability Science*, 8(2), 279–293.
- Mobjörk, M. (2010). Consulting versus participatory transdisciplinarity: A refined classification of transdisciplinary research. *Futures*, 42(8), 866–873.
- Moed, H. F. (2017). *Applied evaluative informetrics*. Cham: Springer.
- OPENing UP. (2016). *Deliverable D5.1—Altmetrics Status Quo. OPENing UP new methods, indicators and tools for peer review, impact measurement and dissemination of research*. <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ae8a151c&appId=PPGMS> Project acronym: OpenUP. Retrieved from Brussels, Belgium.
- Pohl, C. (2010). From transdisciplinarity to transdisciplinary research. *Transdisciplinary Journal of Engineering Science*, 1, 65–73.
- Priem, J. (2014). Altmetrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond Bibliometrics: Harnessing Multi-Dimensional Indicators of Performance*, pp. 263–288. Cambridge, MA: MIT Press.
- Priem, J., & Hemminger, B. M. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday*, 15(7).
- Puschmann, C. (2014). (Micro) blogging science? Notes on potentials and constraints of new forms of scholarly communication. In S. Bartling & S. Friesike (Eds.), *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, pp. 89–106. Cham: Springer.
- Radhakrishna, S. (1965). Combination of results from several 2×2 contingency tables. *Biometrics*, 21, 86–98.
- Ringelhan, S., Wollersheim, J., & Welp, I. M. (2015). I like, I cite? Do Facebook likes predict the impact of scientific work? *PLOS ONE*, 10(8), e0134389. <https://doi.org/10.1371/journal.pone.0134389>
- Robinson-Garcia, N., Costas, R., Isett, K., Melkers, J., & Hicks, D. (2017). The unbearable emptiness of tweeting—About journal articles. *PLOS ONE*, 12(8), e0183551. <https://doi.org/10.1371/journal.pone.0183551>
- Rons, N. (2012). Partition-based field normalization: An approach to highly specialized publication records. *Journal of Informetrics*, 6(1), 1–10. <https://doi.org/10.1016/j.joi.2011.09.008>
- Rons, N. (2014). Investigation of Partition Cells as a Structural Basis Suitable for Assessments of Individual Scientists. In P. Wouters (Ed.), *Proceedings of the Science and Technology Indicators Conference 2014 Leiden “Context Counts: Pathways to Master Big and Little Data”*, pp. 463–472. Leiden, Netherlands: University of Leiden.
- Samoilenko, A., & Yasseri, T. (2014). The distorted mirror of Wikipedia: A quantitative analysis of Wikipedia coverage of academics. *EPJ Data Science*, 3(1), 1.
- SDSN Australia/Pacific. (2017). *Getting started with the SDGs in universities: A guide for universities, higher education institutions, and the academic sector. Australia, New Zealand and Pacific Edition*. Sustainable Development Solutions Network—Australia/Pacific, Melbourne.
- Serrano-López, A. E., Ingwersen, P., & Sanz-Casado, E. (2017). Wind power research in Wikipedia: Does Wikipedia demonstrate direct influence of research publications and can it be used as adequate source in research evaluation? *Scientometrics*, 112(3), 1471–1488. <https://doi.org/10.1007/s11192-017-2447-2>
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2012a). Research blogs and the discussion of scholarly information. *PLOS ONE*, 7(5), e35869. <https://doi.org/10.1371/journal.pone.0035869>
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2012b). Self-citation of bloggers in the science blogosphere. In A. Tokar, M. Beurskens, S. Keuneke, M. Mahrt, I. Peters, C. Puschmann, K. Weller, & T. van Treeck (Eds.), *Science and the Internet*, pp. 183–192. Düsseldorf, Germany: Düsseldorf University Press.
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2014a). Scholarly blogs are a promising altmetric source. *Research Trends*, 37, 11–13.
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2014b). Do blog citations correlate with a higher number of future citations? Research

- blogs as a potential source for alternative metrics. *Journal of the Association for Information Science and Technology*, 65(5), 1018–1027. <https://doi.org/10.1002/asi.23037>
- Sheskin, D. (2007). *Handbook of parametric and nonparametric statistical procedures* (4th Ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2016). Scholarly use of social media and altmetrics: A review of the literature. Retrieved from <https://arxiv.org/abs/1608.08112>
- Tattersall, A., & Carroll, C. (2018). What can Altmeteric.com tell us about policy citations of research? An analysis of Altmeteric.com data for research articles from the University of Sheffield. *Frontiers in Research Metrics and Analytics*, 2(9). <https://doi.org/10.3389/frma.2017.00009>
- Teplitskiy, M., Lu, G., & Duede, E. (2017). Amplifying the impact of open access: Wikipedia and the diffusion of science. *Journal of the Association for Information Science and Technology*, 68, 2116–2127. <https://doi.org/10.1002/asi.23687>
- Thune, T., Reymert, I., Gulbrandsen, M., & Olaf Aamodt, P. (2016). Universities and external engagement activities: Particular profiles for particular universities? *Science and Public Policy*, 43(6), 774–786. <https://doi.org/10.1002/asi.23687>
- Vilkins, S., & Grant, W. J. (2017). Types of evidence cited in Australian Government publications. *Scientometrics*, 113(3), 1681–1695. <https://doi.org/10.1007/s11192-017-2544-2>
- Wiek, A., Talwar, S., O'Shea, M., & Robinson, J. (2014). Toward a methodological scheme for capturing societal effects of participatory sustainability research. *Research Evaluation*, 23(2), 117–132.
- Wooldridge, J., & King, M. B. (2019). Altmeteric scores: An early indicator of research impact. *Journal of the Association for Information Science and Technology*, 70(3), 271–282. <https://doi.org/10.1002/asi.24122>
- Yarime, M., Trencher, G., Mino, T., Scholz, R. W., Olsson, L., Ness, B., ... Rotmans, J. (2012). Establishing sustainability science in higher education institutions: Towards an integration of academic development, institutionalization, and stakeholder collaborations. *Sustainability Science*, 7(1), 101–113.
- Yu, H. (2017). Context of altmetrics data matters: An investigation of count type and user category. *Scientometrics*, 111(1), 267–283. <https://doi.org/10.1007/s11192-017-2251-z>